



## Practice of Epidemiology

### On the Pitfalls of Adjusting for Gestational Age at Birth

Allen J. Wilcox\*, Clarice R. Weinberg, and Olga Basso

\* Correspondence to Dr. Allen Wilcox, Epidemiology Branch (MD A3-05), National Institute of Environmental Health Sciences, P.O. Box 12233, Durham, NC 27709 (e-mail: wilcox@niehs.nih.gov).

*Initially submitted February 23, 2011; accepted for publication June 6, 2011.*

Preterm delivery is a powerful predictor of newborn morbidity and mortality. Such problems are due to not only immaturity but also the pathologic factors (such as infection) that cause early delivery. The understanding of these underlying pathologic factors is incomplete at best. To the extent that unmeasured pathologies triggering preterm delivery also directly harm the fetus, they will confound the association of early delivery with neonatal outcomes. This, in turn, complicates studies of newborn outcomes more generally. When investigators analyze the association of risk factors with neonatal outcomes, adjustment for gestational age as a mediating variable will lead to bias. In the language of directed acyclic graphs, gestational age is a collider. The theoretical basis for colliders has been well described, and gestational age has recently been acknowledged as a possible collider. However, the impact of this problem, as well as its implications for perinatal research, has not been fully appreciated. The authors discuss the evidence for confounding and present simulations to explore how much bias is produced by adjustments for gestational age when estimating direct effects. Under plausible conditions, frank reversal of exposure-outcome associations can occur. When the purpose is causal inference, there are few settings in which adjustment for gestational age can be justified.

adjustment; collider; directed acyclic graph; gestational age; infant mortality; mediating variable; premature birth; stratification

Abbreviation: OR, odds ratio.

Recent developments in clinical medicine and epidemiologic theory have provided important insights about the links between preterm birth and infant health. There are heterogeneous (and mostly unrecorded) pathologies that cause early delivery and also directly harm the fetus (1). Such unmeasured pathologies amplify the apparent effect of immaturity on mortality or morbidity (2). Furthermore, these pathologic conditions distort causal inferences that rely on adjustments for gestational age as an intermediate (or mediating) variable. In the terms of directed acyclic graphs, gestational age is a “collider.” The extent to which routine adjustments for gestational age might bias causal inference has only begun to be explored (3, 4). We describe the evidence for gestational age as a collider and develop simple scenarios to estimate the bias produced by adjustments for gestational age. Our discussion focuses on preterm birth and mortality, but the same arguments apply to gestational age in general and to any morbidity or long-term health outcome associated with gestational age.

#### PRETERM BIRTH AND MORTALITY: THE PROBLEM OF CONFOUNDING

The physiologic development recognized as fetal maturation equips the fetus to survive outside the uterus. The most accessible measure of fetal maturation is gestational age at birth. Infants born before 37 completed weeks of gestation are defined as preterm (or premature).

Immaturity undoubtedly contributes to the higher mortality and morbidity observed among preterm babies. Preterm delivery, however, is itself the result of pathologic processes (5). In addition to immaturity, preterm babies carry the burden of whatever pathology has triggered their early birth. For example, some malformations lead to earlier delivery (6) and, at the same time, independently increase the baby’s risk of death. Severe malformations are usually recorded at birth, but other causes of preterm delivery are often unrecognized or unrecorded: infection of the amniotic or fetal membranes,

disorders of the attachment of the placenta, uteroplacental ischemia, and others yet to be recognized (7–14). Babies born early are smaller on average than fetuses remaining in utero, providing further evidence that preterm babies are affected by pathologic factors before delivery (15).

### The elusive effect of immaturity

Unmeasured factors that cause early delivery and also harm the fetus will strengthen (confound) the observed association between early delivery and newborn mortality (2). The mortality caused solely by immaturity is thus unobservable. An indirect attempt to estimate this risk suggests that immaturity contributes only half of preterm mortality, perhaps even less (16).

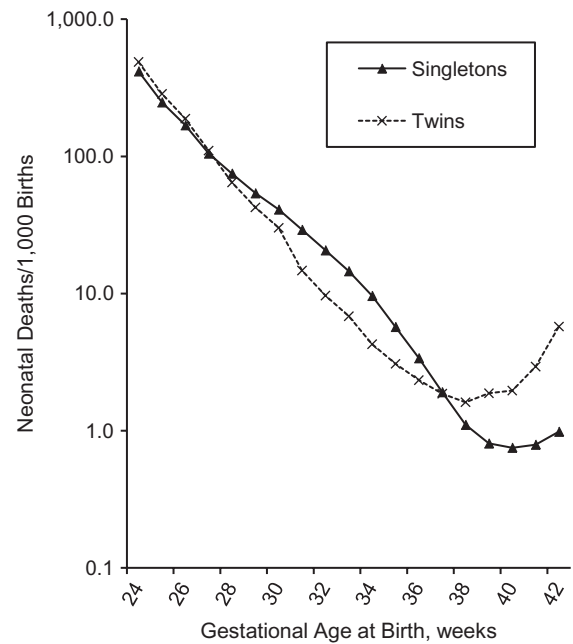
It is a useful thought experiment to consider a randomized trial in which women are assigned to deliver early. To the extent that pathologic factors cause preterm births and independently add risk, the babies of women who deliver by spontaneous onset of labor would have worse survival than the babies of women randomly selected for preterm delivery at the same gestational age. Support for this hypothesis is suggested by a recent analysis of late-preterm deliveries by women who had been induced with “no medical indication” (17). Babies born of induced deliveries were at markedly lower risk of virtually every indicator of morbidity compared with babies born spontaneously at the same gestational age—evidence that early spontaneous labor and delivery produce babies with an additional burden of pathology.

### Heterogeneity in the causes of preterm birth

Some preterm births at a given gestational age are worse off than others. This heterogeneity could arise for at least 2 reasons. First, the pathologic factors that cause preterm delivery presumably vary in the strength of their effect on newborn mortality. Second, the pathologic causes of early delivery seem to vary in their effects on the duration of pregnancy. For example, acute chorioamnionitis tends to be associated with early preterm delivery, while chronic chorioamnionitis is found more often in later preterm delivery (13). Specific placental vascular lesions are more common before 34 weeks than at 35–36 weeks (9). These variations produce an uneven distribution of causes of preterm delivery across gestational age.

### Heterogeneity and the intersection of mortality curves

Heterogeneity of causes of preterm birth plausibly explains the patterns of mortality seen with known risk factors for preterm delivery. For example, one known risk factor is twinning, a natural variant of human reproduction thought to trigger early delivery through overcrowding of the uterus (and, in the modern era, through medical intervention). Although twinning carries increased mortality risk overall (18), a preterm twin would probably be better off than a singleton delivered early because of (for example) a severe malformation. More generally, twins at a given preterm gestational age may be healthier as a group than singletons, who, by definition, delivered prematurely for other reasons, some of them quite dangerous.



**Figure 1.** Week-specific neonatal mortality of singletons and twins among US livebirths for gestational-age weeks 24–42, National Center for Health Statistics, 1995–2002. Gestational age is based on last menstrual period, replaced by clinical gestation when last menstrual period was implausible (*Epidemiology* 2010;21(4):521–527 (16)).

A twin advantage during the preterm period is in fact seen in the gestational-age-specific differences in mortality for twins and singletons (Figure 1). Before 28 weeks, twins and singletons have similarly high mortality, suggesting that twin and singleton babies born so extremely early may carry similar burdens of pathology and immaturity (or perhaps that extreme immaturity swamps other effects). Starting at 28 weeks, preterm twins have markedly better survival than preterm singletons. Assuming that singletons and twins are equally immature at a given early gestational age, the mortality advantage of twins suggests that preterm twins have fewer underlying pathologies than preterm singletons (19). This is supported by clinical research: Twins at 28–36 weeks have fewer intracerebral hemorrhages, lower risk of neonatal septicemia, and fewer birth defects than singletons at the same preterm ages (20).

After 37 weeks, twins have higher mortality. This may be because singletons born at term comprise a less selected (and presumably healthier) sample than singletons born preterm, thus allowing the higher risk of twinning to become apparent.

Another explanation sometimes advanced for the pattern in Figure 1 is that twin fetuses born preterm have been stressed in utero, which accelerates their lung maturation and thus provides a physiologic advantage after birth (21). This explanation, however, is not well supported by biologic assessments of human fetal lung maturation (22). Gestational-age-matched comparisons of twins and singletons have reported no differences in markers of lung maturation (23).

Other hypotheses could no doubt be constructed to explain the intersecting mortality curves in Figure 1 but, like fetal lung maturation, they would require a biologic interaction; that is, twinning would have to affect newborn mortality differently at different gestational ages. In contrast, the hypothesis of unmeasured confounding is not only biologically plausible but also parsimonious (2). It could also explain why intersecting gestational-age-specific mortality curves occur across diverse comparisons (e.g., babies born to mothers with pregnancy-induced hypertension compared with those born of normotensive mothers (24) and African-American babies compared with white babies (25)). A rigorous test of the hypothesis of unmeasured confounding would require information on all the pathologic conditions that cause preterm delivery. If confounding were responsible for the intersections of mortality curves, then complete control for those factors would remove the intersections.

An analogous argument has been made to explain the intersection of birth-weight-specific mortality curves (26–30). Low-weight babies from high-risk groups consistently have lower mortality than low-weight babies from low-risk groups. Unmeasured factors are likely to confound the association of birth weight with infant mortality, causing birth weight to become a collider (27–29). In turn, the construction of mortality curves across strata of the collider (birth weight) leads to intersections and reversals of risk (29, 30).

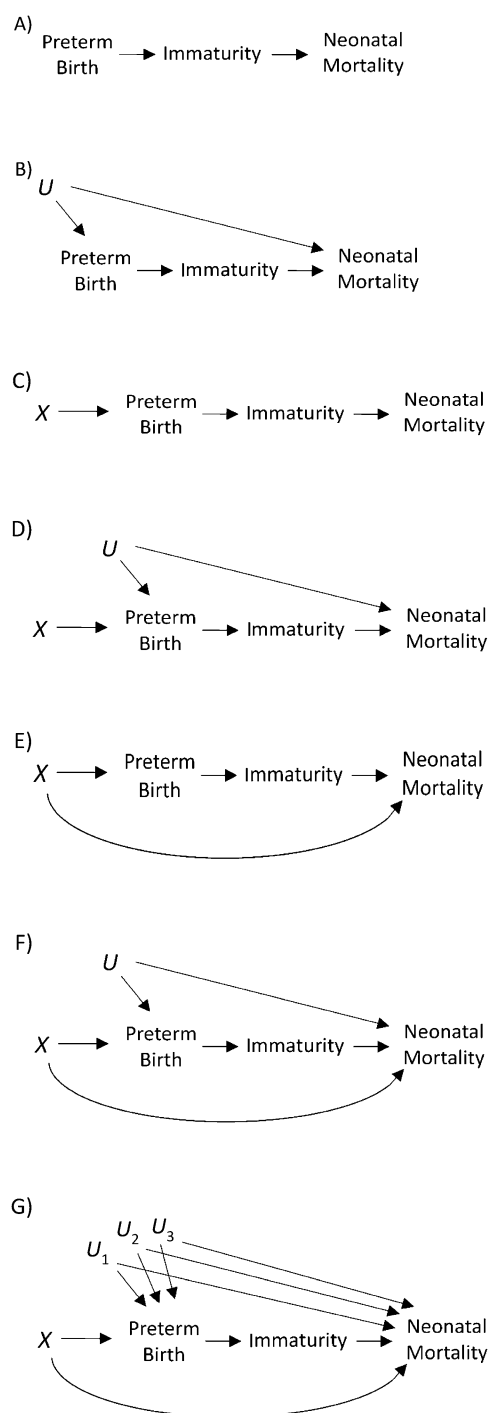
Researchers have been quicker to grasp the possibility that birth weight is a collider than the possibility that gestational age is a collider. This may be because early delivery unquestionably contributes to mortality, while low weight does not necessarily. Premature birth evicts an infant who is not fully prepared to cope with extrauterine life, whereas there is no physiologic reason that small size (within a reasonable range) should increase a baby's risk. The possibility of unmeasured confounding may thus be more obvious for low birth weight than for early delivery. Nonetheless, the extent of confounding with gestational age may be extensive.

### PRETERM BIRTH AS A MEDIATING (INTERMEDIATE) VARIABLE

Epidemiologists often seek to assess the direct effect of some prenatal factor (e.g., smoking) on newborn mortality. Because prenatal smoking can cause early delivery (31), gestational age may be regarded as an intermediate on a causal pathway from smoking to newborn mortality. It has long been routine to adjust for the mediating effect of gestational age, even though adjustment for an intermediate variable can produce bias when unmeasured confounders act on both the intermediate and the outcome (32).

Simple directed acyclic graphs provide a useful format for describing these relations (33). The basic association of preterm birth and neonatal mortality is expressed in Figure 2A: Preterm babies suffer a lack of physiologic preparedness for extrauterine life that leads to mortality.

Preterm babies can also carry a mortality risk conferred by whatever condition led to their early delivery (Figure 2B). Such factors (represented by  $U$ ) are poorly characterized in epidemiologic studies and may be unmeasured clinically. In Figure 2B,  $U$  acts as a confounder, in this case adding to the



**Figure 2.** Directed acyclic graphs corresponding to alternative models A–G for the causal association of preterm birth and neonatal mortality. “X” represents a measured variable of interest to the investigator, and “U” represents unmeasured confounding variables.

strength of the association between preterm delivery and mortality.

Now consider ways in which epidemiologists study preterm birth. Preterm delivery is an appealing endpoint because it is

strongly associated with mortality while being much more common than mortality. (Developed countries have 40–175 preterm births per 1,000 livebirths (34) compared with only 2–6 neonatal deaths per 1,000 livebirths (35)). Epidemiologists frequently evaluate perinatal risk factors (such as maternal smoking, exercise, or alcohol use) (31) by their associations with preterm delivery. The usual and reasonable assumption is that risk factor  $X$  that causes earlier delivery will thereby increase the risk of mortality and morbidity (Figure 2C).

A limitation of this interpretation is that the link between preterm delivery and mortality has been strengthened by the presence of  $U$  (Figure 2D). Because the unmeasured factors represented by  $U$  can vary from setting to setting, it is not possible to predict with confidence the extent to which a factor  $X$  that increases (or decreases) the risk of preterm birth will affect mortality and morbidity.

These difficulties extend to studies that seek to assess direct effects of any risk factor  $X$  on neonatal outcomes (Figure 2E). In assessing the association of  $X$  with an outcome, researchers may find it natural (and sometimes even required by reviewers or editors) to adjust for gestational age. Researchers may, for example, wish to know whether an exposure of interest exerts a direct harm on the fetus, separate from the indirect harm inflicted by being born early. Such adjustment is presumed to provide an estimate of the “direct effect” of  $X$  by removing the intermediate influence of early delivery.

However, a direct effect can be estimated only if no uncontrolled confounders are working on preterm birth and mortality, an assumption that is almost certainly untrue (1). Figure 2F shows the more realistic picture, with  $U$  acting as a confounder in the association of preterm birth and mortality. Preterm delivery is then a collider (33). Adjustment for preterm birth blocks the effect of preterm on mortality and creates a backdoor path from  $X$  to mortality via  $U$ . This biases the assessment of the direct effect of  $X$  on mortality. Any adjustment for preterm (or, indeed, any stratification or restriction that takes into account the distribution of gestational age) is likely to bias the estimated direct causal effect of  $X$ .

Epidemiologists and clinicians frequently study the association of preterm birth with morbidity or mortality. Such associations are interpretable as simple predictions, but when etiologic questions are raised (e.g., “why do preterm babies with  $X$  fare better than preterm babies without  $X$ ?”), the conclusions can easily be distorted. This distortion can be thought of as a variant of Berkson’s bias (36), in which patients hospitalized with one disease are compared with other hospitalized patients. The result depends on the specific diseases that have caused the “control” group to be hospitalized, just as a comparison among preterm births depends on the specific diseases that have caused the “control” babies to be born preterm.

## EMPIRICAL ESTIMATES OF BIAS

When faced with a hypothetical bias, epidemiologists must ask the practical question, “How much does it really matter?” In the present context, we can ask, “How strong is the collider bias with adjustment for gestational age under plausible conditions?” We explore this question with calculations based on a simple simulation and using the same parameters previously

used to show that rare, potent factors may be responsible for most of preterm mortality (2). The directed acyclic graph structure of our scenarios is provided in Figure 2G.

Briefly, we assume that, in the absence of any pathology, babies are born at a “target” gestational age, which has a Gaussian distribution in the population (mean of 40 weeks and standard deviation of 10 days) (37). We restrict the model to babies delivered at 24–44 weeks. We impose an arbitrarily defined functional relation between gestational age and mortality to represent the baseline effect of fetal immaturity (the arrows from “preterm delivery” through “immaturity” to “neonatal mortality”). This week-specific “baseline mortality” (i.e., the mortality that each baby would experience from immaturity alone if randomly delivered at any given week) is given by the quadratic function:

$$\ln(p(M)) = a + bz + cz^2,$$

where  $p(M)$  is the probability of infant death;  $z$  is the week of gestation at birth (normalized by subtracting the mean (40 weeks) and dividing by the standard deviation (10 days or 1.43 weeks)); and coefficients  $a$ ,  $b$ , and  $c$  are, respectively,  $-8$ ,  $-0.15$ , and  $0.036$ . In this model, a baby whose target gestation is 37 weeks has a higher mortality than a baby whose target gestation is 40 weeks.

We then add 3 “unmeasured” confounding factors ( $U_1$ ,  $U_2$ , and  $U_3$ ) that cause early delivery and also harm the fetus in other ways. One  $U$  has a relatively weak direct effect on mortality, and the other 2 have strong mortality effects. The weak factor ( $U_1$ ) is present in 4% of the population, reduces gestational age by 35 days, and directly increases mortality with an odds ratio of 1.5. The second factor ( $U_2$ ) is rare (0.6%), reduces gestational age by 50 days, and has a direct mortality odds ratio of 8. The third factor ( $U_3$ ) is the same as the second except that its effect on mortality starts as an odds ratio of 4 at 24 weeks and then increases by 30% with each additional week of gestation. This factor represents a condition that, once established, creates an increasingly hostile intra-uterine environment. As clinicians have recognized, the delay of early labor and delivery may actually increase risk (37). Among those affected by  $U_3$ , the odds ratio is 4 with delivery at week 24, 5.2 with delivery at week 25, 6.76 at week 26, and so on. About 99% of babies with  $U_3$  are born by week 36, at which time the odds ratio among those with  $U_3$  reaches 93. In this model, the 3  $U$ ’s are assumed to act independently of one another (and of  $X$ ) and to act multiplicatively on baseline mortality.

With this overall structure, we now consider the analysis of a measured factor ( $X$ ) that is of interest to the investigator.  $X$  reduces gestational age by 20 days (similar to twinning) and independently increases mortality by a direct-effect odds ratio of 1.7. In our model, the unadjusted (or composite) mortality odds ratio for babies with  $X$  is 2.9. (The unadjusted mortality combines the direct effect of  $X$  on mortality with the more indirect mortality that occurs through the effect of  $X$  on gestational age).

The key question is how well the true direct effect of  $X$  on mortality (odds ratio (OR) = 1.7) can be estimated by adjusting the observed odds ratio of 2.9 for gestational age alone (i.e., by removing the effects of gestational age but not the

effects of the unmeasured  $U$  factors that confound gestational age and mortality). We ran a simulation with 16 million babies, 40% of whom had  $X$ . The delivery date for each baby was determined by its target gestation (or, if the baby was affected by  $X$  or the  $U$ 's, the target gestation minus the number of weeks of early delivery caused by those factors). To avoid residual confounding in the logistic regression, all births were forced to occur at exactly midweek. Despite the large numbers, some minor random variation occurs due to the stochastic nature of the simulation.

With this as our constructed data set, we were able to carry out adjustments for gestational age. Adjusting the effect of  $X$  for gestational age (in 1-week categories) produced a mortality odds ratio of 0.74, in contrast with its true effect of 1.7. Adjustment for gestational age reversed the apparent direction of effect, mistakenly suggesting that  $X$  is protective.

We explored the robustness of this result by performing sensitivity analyses in which parameters of the simulation were substantially altered. Specifically, we changed the underlying risk for immaturity to be linear instead of quadratic ( $\ln(p(M)) = -6 + (-0.15)z$ ); we weakened the "strong" unmeasured confounders by up to half (OR with  $U_2$  reduced from 8 to 5 and the weekly increase in mortality OR with  $U_3$  reduced from 1.30 to 1.15); and we weakened the effect of the main variable ( $X$ ) on gestational age from 20 days to 10 days. These changes were implemented one at a time, with the conditions otherwise held the same as in the main model. We then repeated the adjustment analyses. The adjusted effect of the main variable remained substantially negatively biased in each scenario (0.93, 1.05, and 1.17, respectively, compared with the true value of 1.70).

The strength of this bias is consistent with other recent evidence suggesting the presence of powerful unmeasured confounders (1, 2). It has been shown through a model (the same model on which our simulations were based) that a large portion of preterm mortality could, in theory, be due to relatively rare (although strong) unmeasured confounding factors (2). On an intuitive level, the bias with adjustment can be understood as the consequence of overestimating the true effect of gestational age. When an exaggerated effect of preterm birth is removed through adjustment, the remaining direct effect attributed to the known factor  $X$  is underestimated, even to the point of being reversed.

Our results are based on simplistic (although not unrealistic) assumptions in an arbitrarily quantified causal model. If our scenarios are as plausible as intended, it is possible that similar biases with gestational-age adjustment could be at play in real data. We explored this by estimating the effect of twinning in the US 2002 linked vital-statistics data set, with neonatal mortality as the outcome and gestational age based on last menstrual period (after correcting for major errors) (16). The unadjusted odds ratio for mortality among twins compared with singletons was 5.4. After adjustment for gestational age at birth, the "direct" odds ratio for twins was 0.89.

The reversal of twinning risk after gestational-age adjustment has not (to the best of our knowledge) been discussed in the literature. Others who may have done these calculations perhaps found the results too unlikely (or too confusing) to publish. Maybe they recognized that, in light of the intersecting curves of Figure 1, any summary measure of relative risk

across the gestational-age strata is uninterpretable. Regardless, given the many complications of pregnancy and delivery to which twins are prone (18), it is implausible that twinning could "improve" survival after adjusting for (blocking the effects of) gestational age. If anything, twinning must add to the risk produced by earlier delivery. The "protective" effect of twinning in this example suggests that the bias with gestational-age adjustment is substantial.

Our final exploration was to see how well the partial control of confounding might reduce bias. We repeated our analyses with adjustments for the 3 "unmeasured" confounders, one at a time and in all combinations. Adjustment for all 3  $U$ 's provided a good estimate of the direct effect of  $X$  of 1.72, as we would expect (Table 1). However, when either of the 2 strong confounding factors ( $U_2$  or  $U_3$ ) was left out of the adjustment, there was substantial bias in the estimated direct effect of  $X$  (with OR values ranging from 0.72 to 1.25). In the presence of 2 rare and strong confounding factors, an adjustment for one did surprisingly little to reduce bias in the estimated direct effect. In practical terms, the discouraging message seems to be that, until every strong cause of preterm delivery can be taken into account, adjustments for gestational age are likely to produce biased estimates.

## IMPLICATIONS FOR ETIOLOGIC RESEARCH

As many have previously noted (1, 3, 11, 38), the risk experienced by preterm infants has as much to do with the reasons for being born preterm as with the timing of delivery. To the extent that these pathologies remain unmeasured, they complicate any inference about the risk due to immaturity (4). This problem undermines any causal assessment that depends on gestational-age adjustment. If we wish to estimate the causal effect of a prenatal factor on newborn mortality (or on any other adverse birth outcome), the safest estimate is one that adjusts for confounding factors but not for gestational age.

There are probably few settings in which gestational age itself can be a confounder. Gestational age (or impending delivery) would have to be a cause of the exposure or clinical condition under study (33). One example would be a study of oxygen therapy in newborns and its effects on retinopathy of prematurity. In this case, preterm delivery causes the administration of oxygen and also confers an increased risk of retinopathy, thus serving as a true confounder that must be adjusted for. Outside of such settings, refraining from adjustment is the only sure way to avoid potentially damaging collider bias.

There may be ways in which certain bounds on direct effects can be indirectly inferred through sensitivity analyses (39). The assumptions necessary for such calculations, however, have yet to be fully assessed. Another methodological issue is with regard to the definition of gestational-age-specific risk. A "fetuses-at-risk approach" has been suggested for such analyses (40). However, given that the discussion here is of the risks experienced by those who are already born, the fetuses-at-risk approach seems not to apply, as the effects of immaturity are not a concern for the unborn (41, 42).

As an endpoint in itself, preterm delivery may not be as dependable for investigating prenatal hazards as has been assumed. Two populations might differ in their mix of

**Table 1.** Model Showing Estimates of the Mortality Effect (Expressed as Odds Ratios) of a Known Variable *X* After Adjustments for Gestational Age Alone and With All Possible Combinations of 3 Confounders<sup>a</sup>

Variables Included in the Model	<i>X</i>	<i>U</i> <sub>1</sub>	<i>U</i> <sub>2</sub>	<i>U</i> <sub>3</sub>	( <i>U</i> <sub>3</sub> × <i>GA</i> ) Interaction
<i>Odds Ratios as Estimated by the Model</i>					
None	2.88				
GA	0.74				
GA + <i>U</i> <sub>1</sub>	0.74	0.25			
GA + <i>U</i> <sub>2</sub>	0.77		1.32		
GA + <i>U</i> <sub>3</sub>	0.98			4.58	
GA + <i>U</i> <sub>1</sub> + <i>U</i> <sub>2</sub>	0.72	0.23	0.74		
GA + <i>U</i> <sub>1</sub> + <i>U</i> <sub>3</sub>	0.89	0.39		2.92	
GA + <i>U</i> <sub>2</sub> + <i>U</i> <sub>3</sub>	1.46		8.25	15.43	
GA + <i>U</i> <sub>1</sub> + <i>U</i> <sub>2</sub> + <i>U</i> <sub>3</sub>	1.70	1.78	13.82	25.86	
GA + <i>U</i> <sub>3</sub> + interaction	1.25			0.42	1.57
GA + <i>U</i> <sub>2</sub> + <i>U</i> <sub>3</sub> + interaction	1.55		5.43	2.46	1.33
GA + <i>U</i> <sub>1</sub> + <i>U</i> <sub>3</sub> + interaction	1.14	0.48		0.35	1.52
GA + <i>U</i> <sub>1</sub> + <i>U</i> <sub>2</sub> + <i>U</i> <sub>3</sub> + interaction	1.72	1.53	8.15	4.00	1.31
<i>True Causal Odds Ratio</i>					
	1.70	1.5	8.0	4.0	1.3

Abbreviations: GA, gestational age; *U*<sub>1,2,3</sub>, unknown confounding factors; *X*, known factor of interest.

<sup>a</sup> *X* shortens gestational age by 20 days and carries a true mortality risk odds ratio = 1.7. Three confounders: *U*<sub>1</sub> (population prevalence = 0.04, effect on gestational age –35 days, odds ratio = 1.5); *U*<sub>2</sub> (population prevalence = 0.006, effect on gestational age –50 days, odds ratio = 8); and *U*<sub>3</sub> (population prevalence = 0.006, effect on gestational age –50 days, odds ratio = 4 at 24 weeks, plus a per-week increase of odds ratio = 1.3).

unmeasured factors causing preterm deliveries, such that the 2 groups have the same occurrence of preterm births but with different mortality among the preterm. In such settings, population differences in the incidence of preterm delivery will not necessarily produce predictable differences in newborn mortality. An example would be a population with a large proportion of pregnancies conceived through artificial reproductive techniques, which may have a high risk of preterm due to multiple births. Such a population could be better off than another population with fewer preterm births but a high rate of prenatal infections among the preterm.

As suggested at the outset, the arguments raised here apply to morbidity as well as to mortality. A host of disabilities are associated with preterm delivery, including cerebral palsy, mental retardation, learning and behavioral abnormalities, and psychologic disturbances (43). Immaturity at birth presumably contributes to these health problems, but the pathologic factors that cause preterm delivery are probably important contributors as well (37). The role of gestational age as a predictor of adverse outcomes is in no way challenged by the issues raised here. However, actual effects of newborn immaturity on childhood disability (as on mortality) may be weaker than generally assumed. The same questions can be raised about the causal association of postterm delivery with poor outcomes.

Gestational age at birth is one of the strongest predictors of infant survival. Adjustment for gestational age as a mediating variable has been accepted as routine and even necessary

by many perinatal epidemiologists, including us. To avoid “adjustment” means eschewing virtually every analytical approach in etiologic research that stratifies by gestational age—logistic regression, standardization, matching, restriction, and others. We believe the practical implications of these emerging ideas are only beginning to be appreciated. Although this shift in thinking may be difficult, it points to the crucial distinction—here, as in all areas of epidemiology—between descriptive and causal models. When causation is the question, adjustment for gestational age will seldom provide a trustworthy answer.

## ACKNOWLEDGMENTS

Author affiliations: Epidemiology Branch, National Institute of Environmental Health Sciences, Durham, North Carolina (Allen J. Wilcox); Biostatistics Branch, National Institute of Environmental Health Sciences, Durham, North Carolina (Clarice R. Weinberg); Department of Obstetrics and Gynecology, McGill University, Montreal, Quebec, Canada (Olga Basso); and Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Quebec, Canada (Olga Basso).

This research was supported by the Intramural Research Program of the National Institutes of Health, National Institute of Environmental Health Sciences.

The authors thank Dr. Matt Longnecker and Dr. Donna Baird for useful comments on an earlier version of this paper and Dr. Robert Platt for seminal discussions on this topic.

Conflict of interest: none declared.

## REFERENCES

- Romero R. Prenatal medicine: the child is the father of the man. 1996. *J Matern Fetal Neonatal Med.* 2009;22(8):636–639.
- Basso O, Wilcox AJ. Might rare factors account for most of the mortality of preterm babies? *Epidemiology.* 2011;22(3):320–327.
- VanderWeele TJ, Hernández-Díaz S. Is there a direct effect of pre-eclampsia on cerebral palsy not through preterm birth? *Paediatr Perinat Epidemiol.* 2011;25(2):111–115.
- Ananth CV, VanderWeele TJ. Placental abruption and perinatal mortality with preterm delivery as a mediator: disentangling direct and indirect effects. *Am J Epidemiol.* 2011;174(1):99–108.
- Greenwood C, Yudkin P, Sellers S, et al. Why is there a modifying effect of gestational age on risk factors for cerebral palsy? *Arch Dis Child Fetal Neonatal Ed.* 2005;90(2):F141–F146.
- Purisch SE, DeFranco EA, Muglia LJ, et al. Preterm birth in pregnancies complicated by major congenital malformations: a population-based study. *Am J Obstet Gynecol.* 2008;199(3):287. e1–287.e8.
- Naeye RL. Editorial: causes and consequences of chorioamnionitis. *N Engl J Med.* 1975;293(1):40–41.
- Klebanoff MA, Shiono PH. Top down, bottom up and inside out: reflections on preterm birth. *Paediatr Perinat Epidemiol.* 1995;9(2):125–129.
- Kelly R, Holzman C, Senagore P, et al. Placental vascular pathology findings and pathways to preterm delivery. *Am J Epidemiol.* 2009;170(2):148–158.
- Kim CJ, Romero R, Kusanovic JP, et al. The frequency, clinical significance, and pathological features of chronic chorioamnionitis: a lesion associated with spontaneous preterm birth. *Mod Pathol.* 2010;23(7):1000–1011.
- Iams JD. The epidemiology of preterm birth. *Clin Perinatol.* 2003;30(4):651–664.
- Romero R, Espinoza J, Mazor M. Can endometrial infection/inflammation explain implantation failure, spontaneous abortion, and preterm birth after in vitro fertilization? *Fertil Steril.* 2004;82(4):799–804.
- Romero R, Chaiworapongsa T, Erez O, et al. An imbalance between angiogenic and anti-angiogenic factors precedes fetal death in a subset of patients: results of a longitudinal study. *J Matern Fetal Neonatal Med.* 2010;23(12):1384–1399.
- Chaiworapongsa T, Romero R, Tarca A, et al. A subset of patients destined to develop spontaneous preterm labor has an abnormal angiogenic/anti-angiogenic profile in maternal plasma: evidence in support of pathophysiologic heterogeneity of preterm labor derived from a longitudinal study. *J Matern Fetal Neonatal Med.* 2009;22(12):1122–1139.
- Secher NJ, Kern Hansen P, Thomsen BL, et al. Growth retardation in preterm infants. *Br J Obstet Gynaecol.* 1987;94(2):115–120.
- Basso O, Wilcox A. Mortality risk among preterm babies: immaturity versus underlying pathology. *Epidemiology.* 2010;21(4):521–527.
- Bailit JL, Gregory KD, Reddy UM, et al. Maternal and neonatal outcomes by labor onset type and gestational age. *Am J Obstet Gynecol.* 2010;202(3):245. e1–245.e12.
- Norwitz ER, Edusa V, Park JS. Maternal physiology and complications of multiple pregnancy. *Semin Perinatol.* 2005;29(5):338–348.
- Lie RT. Invited commentary: intersecting perinatal mortality curves by gestational age—are appearances deceiving? *Am J Epidemiol.* 2000;152(12):1117–1119.
- Jacquemyn Y, Martens G, Ruysinck G, et al. A matched cohort comparison of the outcome of twin versus singleton pregnancies in Flanders, Belgium. *Twin Res.* 2003;6(1):7–11.
- Jobe AH. Lung maturation: the survival miracle of very low birth weight infants. *Pediatr Neonatol.* 2010;51(1):7–13.
- Amiel-Tison C, Cabrol D, Denver R, et al. Fetal adaptation to stress. Part I: acceleration of fetal maturation and earlier birth triggered by placental insufficiency in humans. *Early Hum Dev.* 2004;78(1):15–27.
- Winn HN, Romero R, Roberts A, et al. Comparison of fetal lung maturation in preterm singleton and twin pregnancies. *Am J Perinatol.* 1992;9(5-6):326–328.
- Chen XK, Wen SW, Smith G, et al. Pregnancy-induced hypertension is associated with lower infant mortality in preterm singletons. *BJOG.* 2006;113(5):544–551.
- Papiernik E, Alexander GR, Paneth N. Racial differences in pregnancy duration and its implications for perinatal care. *Med Hypotheses.* 1990;33(3):181–186.
- Wilcox AJ. On the importance—and the unimportance—of birthweight. *Int J Epidemiol.* 2001;30(6):1233–1241.
- Basso O, Wilcox AJ. Intersecting birth weight-specific mortality curves: solving the riddle. *Am J Epidemiol.* 2009;169(7):787–797.
- Hernández-Díaz S, Schisterman EF, Hernán MA. The birth weight “paradox” uncovered? *Am J Epidemiol.* 2006;164(11):1115–1120.
- Hernández-Díaz S, Wilcox AJ, Schisterman EF, et al. From causal diagrams to birth weight-specific curves of infant mortality. *Eur J Epidemiol.* 2008;23(3):163–166.
- Wilcox AJ. Commentary: on the paradoxes of birthweight. *Int J Epidemiol.* 2003;32(4):632–633.
- Savitz DA, Murnane P. Behavioral influences on preterm birth: a review. *Epidemiology.* 2010;21(3):291–299.
- Cole SR, Hernán MA. Fallibility in estimating direct effects. *Int J Epidemiol.* 2002;31(1):163–165.
- Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology.* 1999;10(1):37–48.
- Beck S, Wojdyla D, Say L, et al. The worldwide incidence of preterm birth: a systematic review of maternal mortality and morbidity. *Bull World Health Organ.* 2010;88(1):31–38.
- World Health Organization. *Neonatal and Perinatal Mortality: Country, Regional and Global Estimates, 2000.* Geneva, Switzerland: World Health Organization; 2006.
- Berkson J. Limitations of the application of fourfold table analysis to hospital data. *Biometrics.* 1946;2(3):47–53.
- Moster D, Wilcox AJ, Vollset SE, et al. Cerebral palsy among term and postterm births. *JAMA.* 2010;304(9):976–982.
- Ananth CV, Gyamfi C, Jain L. Characterizing risk profiles of infants who are delivered at late preterm gestations: does it matter? *Am J Obstet Gynecol.* 2008;199(4):329–331.
- VanderWeele TJ. Bias formulas for sensitivity analysis for direct and indirect effects. *Epidemiology.* 2010;21(4):540–551.
- Joseph KS. Incidence-based measures of birth, growth restriction, and death can free perinatal epidemiology from erroneous concepts of risk. *J Clin Epidemiol.* 2004;57(9):889–897.
- Paneth N. Invited commentary: the hidden population in perinatal epidemiology. *Am J Epidemiol.* 2008;167(7):793–796.
- Wilcox AJ, Weinberg CR. Invited commentary: analysis of gestational-age-specific mortality—on what biologic foundations? *Am J Epidemiol.* 2004;160(3):213–214.
- Moster D, Lie RT, Markestad T. Long-term medical and social consequences of preterm birth. *N Engl J Med.* 2008;359(3):262–273.